NASA TECHNICAL MEMORANDUM

NASA TM X-1707

NASA TM X-1707

FORTRAN PROGRAM FOR SPLINE FIT CURVE

by Theodore Katsanis

Lewis Research Center

Cleveland, Ohio

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • DECEMBER 1968

NASA TM X-1707

FORTRAN PROGRAM FOR SPLINE FIT CURVE

By Theodore Katsanis

Lewis Research Center
Cleveland, Ohio

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

ABSTRACT

The spline fit curve is a convenient method for fitting a curve through a given set of points. This report describes a FORTRAN IV computer program which will calculate the spline fit curve, with function values, first and second derivatives, and curvature at any desired interpolated points.

FORTRAN PROGRAM FOR SPLINE FIT CURVE

by Theodore Katsanis

Lewis Research Center

SUMMARY

The spline fit curve is a convenient method for fitting a curve through a given set of points. This report describes a FORTRAN IV computer program which will calculate the spline fit curve, with function values, first and second derivatives, and curvature at any desired interpolated points.

INTRODUCTION

If a set of function values corresponding to a set of arguments is given, there are several ways a curve can be fitted through these values so as to approximate the original function with these values. The classical way is by an n^{th} degree polynomial for $n + 1$ points. However, this may not be satisfactory for a large number of points. Least squares approximations are satisfactory for smoothing, but if the function values are accurate, it is not desired to change these values as a least squares fit would. Another technique is to use fewer points as in four-point Lagrangian interpolation. But this does not lead to a smooth curve since the derivatives will not be continuous.

A method that has received much attention lately is the piecewise cubic, with continuous first and second derivatives, commonly referred to as a spline fit curve (ref. 1). The spline fit curve is a mathematical expression for the shape taken by an idealized spline (thin wood or metal strip) passing through the given points. The spline fit curve gives a simple method of determining an approximating analytical curve which can be used in place of the original curve for interpolation, determining first and second derivatives, curvature, or integration.

The spline fit curve is ideally suited for computer use and has been successfully used in a number of computer programs. This report gives a FORTRAN IV program for the purpose of calculating the spline curve and its integral, with interpolated values of the function, first and second derivatives, and curvatures. This program has READ

and WRITE statements for use as an independent program. The program can also be used as a subroutine with the input and output variables as subroutine arguments.

METHOD AND THEORY

Suppose that an interval $a \leq x \leq b$ is divided into N intervals by $a = x_0 < x_1 < \ldots < x_N = b$. The function F(x) is given at these N + 1 values of x by $F_k = F(x_k)$. Join each given point by a cubic polynomial and require that the first and second derivatives be continuous at these points. Each cubic has four coefficients for a total of 4N unknown coefficients. Each cubic is required to pass through the given points providing 2N conditions to be satisfied. The requirement of continuous first and second derivatives at the N - 1 interior points provides another 2N - 2 conditions, for a total of 4N - 2 conditions to be satisfied. Two additional conditions are necessary to determine the cubic spline curve uniquely. These are usually given as arbitrary end conditions. The end condition used in this program is that the second derivative at an end point is one-half the second derivative at the next point.

The theoretical reason for using a cubic spline curve is that the cubic spline has the minimum value for

$$\int \left[F''(x)\right]^2 dx$$

of all curves passing through the spline points. This is proven in ref. 1. Since

$$\int \left[F''(x)\right]^2 dx$$

is nearly proportional to the strain energy of a thin, uniform spline with a small slope, the cubic spline fit will approximate the shape taken by a thin physical spline passing through the given points. The end condition of making the second derivative at the end point one-half the second derivative at the second point is equivalent to bending the spline beyond the last point slightly, instead of just allowing it to be straight.

In reference 1 there is an extensive discussion of the mathematical theory of spline fit curves. Equation (4) of reference 1 is the basic equation to be solved. A complete set of equations is obtained by specifying values for the parameters λ_0 and λ_N of reference 1 (following eq. (4)). The values of λ_0 and λ_N are determined by the end conditions. The end conditions mentioned above result in $\lambda_0 = \lambda_N = 1/2$. The resulting complete set of equations can be expressed as a tridiagonal matrix equation which is

strongly diagonally dominant. The equation is solved by Gauss elimination. The equation of the spline curve is then equation (3) of reference 1, with the derivative given by equation (2) and the second derivative by equation (1) of reference 1. The curvature K is calculated by

$$K = \frac{F''}{\left[1 + (F')^2\right]^{3/2}}$$

The radius of curvature is the reciprocal of the curvature.

USE OF PROGRAM

The spline fit curve can give very accurate values for first and second derivatives. To do this, however, the spline points must be accurately given. The derivatives are extremely sensitive to small errors in coordinates when points are close together. For this reason it is best to use as few points as is reasonable. If the curvature is slight, the spline points can be widely spaced. The greater the curvature, the closer the spline points should be (analogous to a physical spline). With a very sharp bend, the points should be quite close together and must be very accurate.

If the curve has a large slope, the spline approximation is not as good. There is no trouble if the slope is less than one in absolute value. Much higher slopes can be handled satisfactorily, but more points are required. Great care must be used when using curves approaching the vertical within a few degrees, especially if there is appreciable curvature.

Curves Given by Parametric Equations

The spline curve fit is limited to curves of the form $y = f(x)$. Thus, a circle could not be approximated by a spline fit curve. However, a curve of any shape can be expressed parametrically (ref. 2). For example, x and y can both be given as a function of arc length s. That is,

$$x = x(s)$$

$$y = y(s)$$

For example, for a circle of radius r with center at the origin, we have

$$x = r\cos\left(\frac{s}{r}\right)$$

$$y = r\sin\left(\frac{s}{r}\right)$$

Then the derivative dy/dx can be calculated by

$$\frac{dy}{dx} = \frac{\frac{dy}{ds}}{\frac{dx}{ds}}$$

The second derivative d^2y/dx^2 can be calculated by

$$\frac{d^2y}{dx^2} = \frac{\frac{dx}{ds} \cdot \frac{d^2y}{ds^2} - \frac{dy}{ds} \cdot \frac{d^2x}{ds^2}}{\left(\frac{dx}{ds}\right)^3}$$

The calculation of the first and second derivatives for parametric curves has not been included in the program.

Use as a Subroutine

This program can be used as a subroutine. Most likely, the READ and WRITE statements would be eliminated. The spline input variables (X, Y, and N) would be input arguments as well as the interpolation data (Z and MAX) if interpolation is done. The desired output variables would be the output arguments. Of course, calculation of unnecessary variables could be eliminated.

DESCRIPTION OF INPUT AND OUTPUT

Input. - The input data sheet is shown in figure 1. The quantities filled in are for a sample case. The output for this case is presented in the next section. N and MAX are both integers (no decimal point) and must be right-adjusted. The remaining input is real numbers (punch decimal point) in a 10-column field.

1 5 | 6 10	11 20	21 30	31 40	41 50	51 60	61 70	71 80
TITLE CARD							
Sine curve							
N | MAX							
7 | 8							
X-COORDINATES (SPLINE POINTS)							
0.	0.5235988	1.04719760	1.57079630	2.09439510	2.61799389	3.14159271	
Y-COORDINATES (SPLINE POINTS)							
0.	.5	.86602540	1.	.86602540	.5	0.	
X-COORDINATES (INTERPOLATION POINTS)							
0.	.5	1.	1.5707963	2.	3.	3.1415927	.5235988

Figure 1. - Input form.

The input variables are as follows:

N — Number of spline points (minimum 2, maximum 50)

MAX — Number of interpolation points (minimum 0, maximum 50)

X - COORDINATES (SPLINE POINTS) — Set of X coordinates of the N spline points

Y - COORDINATES (SPLINES POINTS) — Set of Y coordinates of the N spline points

X - COORDINATES (INTERPOLATION POINTS) — Set of X coordinates where output data will be given

Output. - An example of the output is given in figure 2. The first part lists the

SINE CURVE

LIST OF SPLINE COORDINATES (INPUT) (N = 7)

X	Y	VALUE OF THE INTEGRAL OF Y*DX
0	0	0
0.52359880	0.50000000	0.13494445
1.04719760	0.86602540	0.50056853
1.57079630	1.00000000	1.00056283
2.09439510	0.86602540	1.50065725
2.61799389	0.50000000	1.86638130
3.14159271	0	2.00132576

COORDINATES, DERIVATIVES AND CURVATURES AT SELECTED POINTS (NO. OF POINTS = 8)

X(INPUT)	Y(INTERPOLATED)	DY/DX	D2Y/DX2	CURVATURE	RADIUS OF CURVATURE
0	0	1.03361517	-0.22541757	-0.75778930E-01	-13.1962801
0.50000000	0.47956131	0.85709192	-0.44057548	-0.19005273	-5.25159755
1.00000000	0.84133866	0.54372633	-0.86253815	-0.58486097	-1.70980304
1.57079630	1.00000000	0.60070305E-07	-1.01437925	-1.01437925	-0.98582453
2.00000000	0.90936258	-0.41583957	-0.92334693	-0.72687697	-1.37574352
3.00000000	0.14386904	-0.99738205	-0.28637536	-0.10154735	-9.83793373
3.14159271	0	-1.03361511	-0.22541749	-0.75778910E-01	-13.1962837
0.52359880	0.50000000	0.85557263	-0.45083514	-0.19749248	-5.05343377

Figure 2. - Sample output.

spline points given as input, together with

$$\int_{X(1)}^{X(I)} Y \, dX$$

for I = 1, 2, . . . , N.

The next output lists the X coordinates given as interpolation points in the input, with the interpolated values of Y, the first and second derivatives, the curvature, and the radius of curvature.

The message OUT OF RANGE X = ----- is printed if there is extrapolation of more than one-tenth of either of the end intervals. Extrapolated values must be used with caution.

COMPUTER PROGRAM

Program Listing

```
      DIMENSION X(50),Y(50),S(50),A(50),B(50),C(50),F(50),W(50),SB(50),
     1G(50),EM(50),Z(50),YINT(50),DYDX(50),D2YDX(50),CURV(50),RAD(50),
     1SUM(50)
    1 READ (5,1000)
      WRITE(6,1050)
      WRITE(6,1060)
      READ (5,1010) N,MAX
      READ (5,1020)(X(I),I=1,N)
      READ (5,1020)(Y(I),I=1,N)
      READ (5,1020)(Z(I),I=1,MAX)
      DO 10 I=2,N
   10 S(I)=X(I)-X(I-1)
      NO=N-1
      IF(2.GT.NO) GO TO 25
      DO 20 I=2,NO
      A(I)=S(I)/6.0
      B(I)=(S(I)+S(I+1))/3.0
      C(I)=S(I+1)/6.0
   20 F(I)=(Y(I+1)-Y(I))/S(I+1)-(Y(I)-Y(I-1))/S(I)
   25 A(N)=-.5
      B(1)=1.0
      B(N)=1.0
      C(1)=-.5
      F(1)=0.0
      F(N)=0.0
      W(1)=B(1)
      SB(1)=C(1)/W(1)
      G(1)=0.0
      DO 30 I=2,N
      W(I)=B(I)-A(I)*SB(I-1)
      SB(I)=C(I)/W(I)
   30 G(I)=(F(I)-A(I)*G(I-1))/W(I)
```

```
      EM(N)=G(N)
      DO 40 I=2,N
      K=N+1-I
   40 EM(K)=G(K)-SB(K)*EM(K+1)
      SUM(1) = 0.
      DO 45 I=2,N
   45 SUM(I) =SUM(I-1)+S(I)*(Y(I)+Y(I-1))/2.-S(I)**3*(EM(I)+EM(I-1))/24.
      WRITE (6,1030) N,(X(I),Y(I),SUM(I),I=1,N)
      IF(MAX.LT.1) GO TO 1
      DO 90 I=1,MAX
      K=2
      IF(Z(I)-X(1)) 60,50,70
   50 YINT(I)=Y(1)
      GOTO 86
   60 IF(Z(I).LT.(1.1*X(1)-.1*X(2)))WRITE (6,1000)Z(I)
      GO TO 85
   65 IF(Z(I).GT.(1.1*X(N)-.1*X(N-1))) WRITE (6,1000)Z(I)
      K=N
      GO TO 85
   70 IF(Z(I)-X(K)) 85,75,80
   75 YINT(I)=Y(K)
      GOTO 86
   80 K=K+1
      IF(K-N) 70,70,65
   85 YINT(I) = EM(K-1)*(X(K)-Z(I))**3/6./S(K)+EM(K)*(Z(I)-X(K-1))**3/6.
     1/S(K)+(Y(K)/S(K)-EM(K)*S(K)/6.)*(Z(I)-X(K-1))+(Y(K-1)/S(K)-EM(K-1)
     2*S(K)/6.)*(X(K)-Z(I))
   86 DYDX(I)=-EM(K-1)*(X(K)-Z(I))**2/2.0/S(K)+EM(K)*(X(K-1)-Z(I))**2/2.
     10/S(K)+(Y(K)-Y(K-1))/S(K)-(EM(K)-EM(K-1))*S(K)/6.0
      D2YDX(I) = EM(K-1)*(X(K)-Z(I))/S(K)+EM(K)*(Z(I)-X(K-1))/S(K)
      CURV(I) = D2YDX(I)/(1.+DYDX(I)**2)**1.5
      RAD(I) = 1./CURV(I)
   90 CONTINUE
      WRITE (6,1040) MAX,(Z(I),YINT(I),DYDX(I),D2YDX(I),CURV(I),RAD(I),
     1   I=1,MAX)
      GO TO 1
 1000 FORMAT (17HLOUT OF RANGE X =G14.6)
 1010 FORMAT (10I5)
 1020 FORMAT (8G10.8)
 1030 FORMAT (49HL        LIST OF SPLINE COORDINATES (INPUT)    (N = ,I2,
     1   1H),10X,12HVALUE OF THE/15X,1HX,19X,1HY,25X,16HINTEGRAL OF Y*DX
     2   /(5X,2G20.8,G32.8))
 1040 FORMAT (1H1,5X,76HCOORDINATES, DERIVATIVES AND CURVATURES AT SELEC
     1TED POINTS (NO. OF POINTS = ,I2,1H)/12X,8HX(INPUT),8X,15HY(INTERPO
     2LATED),10X,5HDY/DX,14X,7HD2Y/DX2,12X,9HCURVATURE,6X,19HRADIUS OF C
     3URVATURE/(5X,6G20.8))
 1050 FORMAT (1H1)
 1060 FORMAT (80H
     1                                          )
      END
```

FORTRAN Dictionary

A	Array of below diagonal terms of coefficient matrix
B	Array of diagonal terms of coefficient matrix
C	Array of diagonal terms of coefficient matrix
CURV	Array of curvatures at interpolation points
DYDX	Array of derivatives at interpolation points
D2YDX	Array of second derivatives at interpolation points
EM	Array of second derivatives at spline points
F	Array of constants in matrix equation
G	Temporary array used in solving matrix equation
I	DO index
K	DO index
MAX	Number of interpolation points
N	Number of spline points
NO	N - 1
RAD	Array of radii of curvature at spline points
S	Array of length of intervals between spline points
SB	Temporary array used in solving matrix equation
W	Temporary array used in solving matrix equation
X	Array of x-coordinates of spline points
Y	Array of y-coordinates of spline points
YINT	Array of y-coordinates of interpolation points
Z	Array of x-coordinates of interpolation points

Lewis Research Center,
National Aeronautics and Space Administration,
Cleveland, Ohio, September 25, 1968,
126-15-02-31-22.

REFERENCES

1. Walsh, J. L.; Ahlberg, J. H.; and Nilson, E. N.: Best Approximation Properties of the Spline Fit. J. Math. Mech., vol. 11, no. 2, 1962, pp. 225-234.

2. Britton, Jack R.: Calculus. Holt, Rinehart & Winston, Inc., 1956, pp. 63-65.

NASA-Langley, 1968 —— 19 E-4643

"The aeronautical and space activities of the United States shall be conducted so as to contribute . . . to the expansion of human knowledge of phenomena in the atmosphere and space. The Administration shall provide for the widest practicable and appropriate dissemination of information concerning its activities and the results thereof."

— NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and technical information considered important, complete, and a lasting contribution to existing knowledge.

TECHNICAL NOTES: Information less broad in scope but nevertheless of importance as a contribution to existing knowledge.

TECHNICAL MEMORANDUMS: Information receiving limited distribution because of preliminary data, security classification, or other reasons.

CONTRACTOR REPORTS: Scientific and technical information generated under a NASA contract or grant and considered an important contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information published in a foreign language considered to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information derived from or of value to NASA activities. Publications include conference proceedings, monographs, data compilations, handbooks, sourcebooks, and special bibliographies.

TECHNOLOGY UTILIZATION PUBLICATIONS: Information on technology used by NASA that may be of particular interest in commercial and other non-aerospace applications. Publications include Tech Briefs, Technology Utilization Reports and Notes, and Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION DIVISION

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

Washington, D.C. 20546